



Algorithms for some classes of motif finding problems

著者	王 ?
発行年	2015
その他のタイトル	いく種類かのモチーフ探索問題に対するアルゴリズム
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2014
報告番号	12102甲第7269号
URL	http://hdl.handle.net/2241/00125867

氏 名 (本籍)	王 珣
学 位 の 種 類	博 士 (工学)
学 位 記 番 号	博 甲 第 7269 号
学位授与年月日	平成 27 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当
審 査 研 究 科	システム情報工学研究科
学位論文題目	Algorithms for some classes of motif finding problems (いく種類かのモチーフ探索問題に対するアルゴリズム)
主 査	筑波大学 教授 Ph.D. in Combinatorics and Optimization 藤原 良叔
副 査	筑波大学 教授 博士 (工学) 張 勇兵
副 査	筑波大学 教授 工学博士 吉瀬 章子
副 査	筑波大学 教授 博士 (理学) 繁野 麻衣子
副 査	筑波大学 教授 博士 (理学) 繆 瑩
副 査	筑波大学 教授 博士 (工学) イリチュ 美佳
副 査	名古屋大学 教授 理学博士 神保 雅一

論 文 の 要 旨

モチーフとは DNA 配列の複数箇所に潜む比較的短い特定の塩基列をいう。本論文のテーマは与えられた DNA 配列の中からモチーフを発見するアルゴリズムを開発することである。モチーフの塩基は何らかの原因で微妙に変化するものと仮定する。塩基の変化は他のものになってしまう「置換」(substitution)、なくなってしまう「消失」(deletion)、あるいは新しい塩基が加わる「挿入」(insertion)などが考えられる。

第 1 章では、モチーフ発見問題が分子生物学の重要な研究テーマであることを述べ、モチーフ発見問題のモチーフ・モデルを紹介し、モチーフ発見問題のこれまでの研究手法や研究結果を概観している。

第 2 章において、モチーフ塩基の置換のみが起こると仮定した planted モチーフ発見問題を説明し、planted モチーフ発見問題を解くための既存の random projection algorithm とその改良版である uniform projection algorithm を再考している。

第 3 章では、上述 projection algorithm がさらに改善できることを指摘し、low-dispersion sequence にもとづいた low-dispersion projection algorithm を提案した。さらに、low-dispersion

sequence の構成のために、組合せ理論的概念である developed almost difference family を導入した。そして組合せデザイン理論上の手法を駆使し developed almost difference family を数多く構成した。シミュレーションによって low-dispersion projection algorithm の性能が既存の projection algorithm より優れていることを確かめた。

第4章では、この章ではモチーフ塩基の変化は置換、消失、挿入の全てを想定したモチーフ・モデルである planted edited モチーフ発見問題について説明している。そして planted edited モチーフ発見問題を解くための既存の HIGEDA (hierarchical gene-set genetic based algorithm) を評価している。但しこの HIGEDA では置換と消失のみを仮定している。Planted edited モチーフ発見問題は難しい問題であるため、ほとんどの既存の発見アルゴリズムの性能はあまり良くない。

第5章では、置換、消失、挿入の全ての塩基変化を仮定した planted edited モチーフ発見問題を考えている。その発見方法については遺伝アルゴリズムを基本として、期待値最大化法を利用して、3種類のモチーフ塩基変化に対処できるモチーフ発見アルゴリズム GAEM(expectation maximized genetic algorithm) を提案した。そしてシミュレーション実験によって GAEM の性能が HIGEDA より優れていることを確かめた。さらに、ベンチマークテストのために公開されている DNA 配列に対して GAEM を適用することによって、planted edited モチーフ発見問題に関しては、GAEM が代表的な MEME や GLAM2、HIGEDA より全体的に性能が良いことを示した。

第6章において、本研究の結果全般に対する考察と今後の研究課題の展望が示されている。

審 査 の 要 旨

【批評】

モチーフの塩基が微妙に変化する仮定のもとでのモチーフ発見問題において、既存の random projection algorithm と uniform projection algorithm をもとに、組合せデザイン理論の概念を導入した low-dispersion projection algorithm を提案した。つぎに塩基の変化をより現実に近いモデルを想定した planted edited モチーフ発見問題に取り組んだ。この問題では遺伝アルゴリズムを利用した。提案アルゴリズムは全てシミュレーション実験やベンチマークテストによって性能が良くなっていることを示している。これらの研究成果はこの分野に大きく貢献し、博士論文としての水準に達している。

【最終試験の結果】

平成 27 年 1 月 22 日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。